



SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Caras, Ingrid W

(ii) TITLE OF INVENTION: A2-1 Neurotrophic Factor

5 (iii) NUMBER OF SEQUENCES: 8

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: Genentech, Inc.
- (B) STREET: 460 Point San Bruno Blvd
- (C) CITY: South San Francisco
- 10 (D) STATE: California
- (E) COUNTRY: USA
- (F) ZIP: 94080

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
- 15 (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: WinPatin (Genentech)

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER: 08/635130
- 20 (B) FILING DATE: 19-Mar-1996
- (C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Torchia, PhD., Timothy E.
- (B) REGISTRATION NUMBER: 36,700
- 25 (C) REFERENCE/DOCKET NUMBER: P1001

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 415/225-8674
- (B) TELEFAX: 415/952-9881
- (C) TELEX: 910/371-7168

30 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1877 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Double
- 35 (D) TOPOLOGY: Linear

(ix) FEATURE:

- (A) NAME/KEY: Extra Cellular Domain
- (B) LOCATION: 244-899
- (C) IDENTIFICATION METHOD:
- 40 (D) OTHER INFORMATION:

Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg Arg Cys
 80 85 90

GAG GCA CCC CCT GCC CCA AAC CTC CTT CTC ACT TGT GAT 558
 Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
 5 95 100 105

CGC CCA GAC CTG GAT CTC CGC TTC ACC ATC AAG TTC CAG 597
 Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln
 110 115

GAG TAT AGC CCT AAT CTC TGG GGC CAC GAG TTC CGC TCG 636
 10 Glu Tyr Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser
 120 125 130

CAC CAC GAT TAC TAC ATC ATT GCC ACA TCG GAT GGG ACC 675
 His His Asp Tyr Tyr Ile Ile Ala Thr Ser Asp Gly Thr
 135 140

15 CGG GAG GGC CTG GAG AGC CTG CAG GGA GGT GTG TGC CTA 714
 Arg Glu Gly Leu Glu Ser Leu Gln Gly Gly Val Cys Leu
 145 150 155

ACC AGA GGC ATG AAG GTG CTT CTC CGA GTG GGA CAA AGT 753
 Thr Arg Gly Met Lys Val Leu Leu Arg Val Gly Gln Ser
 20 160 165 170

CCC CGA GGA GGG GCT GTC CCC CGA AAA CCT GTG TCT GAA 792
 Pro Arg Gly Gly Ala Val Pro Arg Lys Pro Val Ser Glu
 175 180

ATG CCC ATG GAA AGA GAC CGA GGG GCA GCC CAC AGC CTG 831
 25 Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser Leu
 185 190 195

GAG CCT GGG AAG GAG AAC CTG CCA GGT GAC CCC ACC AGC 870
 Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser
 200 205

30 AAT GCA ACC TCC CGG GGT GCT GAA GGC CCC CTG CCC CCT 909
 Asn Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro
 210 215 220

CCC AGC ATG CCT GCA GTG GCT GGG GCA GCA GGG GGG CTG 948
 Pro Ser Met Pro Ala Val Ala Gly Ala Ala Gly Gly Leu
 35 225 230 235

GCG CTG CTC TTG CTG GGC GTG GCA GGG GCT GGG GGT GCC 987
 Ala Leu Leu Leu Leu Gly Val Ala Gly Ala Gly Gly Ala
 240 245

ATG TGT TGG CGG AGA CGG CGG GCC AAG CCT TCG GAG AGT 1026

| | | |
|----|--|-----|
| | Met Cys Trp Arg Arg Arg Arg Ala Lys Pro Ser Glu Ser | |
| | 250 | 255 |
| | CGC CAC CCT GGT CCT GGC TCC TTC GGG AGG GGA GGG TCT 1065 | |
| | Arg His Pro Gly Pro Gly Ser Phe Gly Arg Gly Gly Ser | |
| 5 | 265 | 270 |
| | CTG GGC CTG GGG GGT GGA GGT ATG GGA CCT CGG GAG 1104 | |
| | Leu Gly Leu Gly Gly Gly Met Gly Pro Arg Glu | |
| | 275 | 280 |
| | 285 | |
| | GCT GAG CCT GGG GAG CTA GGG ATA GCT CTG CGG GGT GGC 1143 | |
| 10 | Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly | |
| | 290 | 295 |
| | 300 | |
| | Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys | |
| | 305 | 310 |
| 15 | GTG AGT GGT GAC TAT GGG CAT CCT GTG TAT ATC GTG CAG 1221 | |
| | Val Ser Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln | |
| | 315 | 320 |
| | 325 | |
| | GAT GGG CCC CCC CAG AGC CCT CCA AAC ATC TAC TAC ACA 1260 | |
| | Asp Gly Pro Pro Gln Ser Pro Pro Asn Ile Tyr Tyr Thr | |
| 20 | 330 | 335 |
| | TCG ATT TCT GTG TTG GAG TGG CCC ATA TTG CAT ACG ATA 1299 | |
| | Ser Ile Ser Val Leu Glu Trp Pro Ile Leu His Thr Ile | |
| | 340 | 345 |
| | 350 | |
| | CAA CTG TTT TTC ATG CGA TCC AAG TGC TCC CGT GTC ACT 1338 | |
| 25 | Gln Leu Phe Phe Met Arg Ser Lys Cys Ser Arg Val Thr | |
| | 355 | 360 |
| | 365 | |
| | ACA TTC TTA TTT CCT GTG CAA GTT ATT ACG ACA TCG ACT 1377 | |
| | Thr Phe Leu Phe Pro Val Gln Val Ile Thr Thr Ser Thr | |
| | 370 | 375 |
| 30 | TGC CGG ATG ACT TCA TTT AGC TTT ACC ACC CTG AAC CCA 1416 | |
| | Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn Pro | |
| | 380 | 385 |
| | 390 | |
| | TCC ATG CAG GCC TGC AGA GCA CAG ATG GGG GAA TTC CGA 1455 | |
| | Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg | |
| 35 | 395 | 400 |
| | ATC AGA TGG TGT TTC TGG GGG GAC AGG ATC CTG GGT ACG 1494 | |
| | Ile Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr | |
| | 405 | 410 |
| | 415 | |
| | GCT CTG TTT GTG CTT GTG ATT CTT CTT CTT GGG AGG 1533 | |

Ala Leu Phe Val Leu Val Leu Ile Leu Leu Leu Gly Arg
420 425 430

CTG AAT ATG CAT CAG ACG ACA CTG CTC CGG CAA CGG GCC 1572
Leu Asn Met His Gln Thr Thr Leu Leu Arg Gln Arg Ala
5 435 440

AGT GTG GAG GCG GAA GCC GGC CAG CAT GGT CCC CTG TG 1610
Ser Val Glu Ala Glu Ala Gly Gln His Gly Pro Leu
445 450 455

ATAGGATTGA AAGAGCTACT GAGAATAGGG GGCTTCTCAA TGAGAGAGCG 1660

10 GAGGCTGCTG TTATCATGGG AACCAAGGCAG ATCAATCATC CCTGGCAGGT 1710
CAGGCAGGAA GTTACTTAGC TTCTCCTTCA CCTTCTTCCC ACAGAATTAA 1760
TTATAGGCTT GTTCCAAGTT GTAGTGTGTG ATCAGATTG TGCTGCCTGT 1810
CAGCTCTGTG CTACCTGGCA GTTCCCCCTCA TGGAAATTGCA TATCAAGCTT 1860
ATCGATAACCG TCGACCT 1877

15 (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 455 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
1 5 10 15

Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
20 25 30

25 Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
35 40 45

Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
50 55 60

Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn
30 65 70 75

Tyr Glu Phe Tyr Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg
80 85 90

Arg Cys Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
95 100 105

Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln Glu Tyr
110 115 120

Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser His His Asp Tyr
125 130 135

5 Tyr Ile Ile Ala Thr Ser Asp Gly Thr Arg Glu Gly Leu Glu Ser
140 145 150

Leu Gln Gly Gly Val Cys Leu Thr Arg Gly Met Lys Val Leu Leu
155 160 165

Arg Val Gly Gln Ser Pro Arg Gly Gly Ala Val Pro Arg Lys Pro
10 170 175 180

Val Ser Glu Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser
185 190 195

Leu Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser Asn
200 205 210

15 Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro Pro Ser Met
215 220 225

Pro Ala Val Ala Gly Ala Ala Gly Gly Leu Ala Leu Leu Leu Leu
230 235 240

Gly Val Ala Gly Ala Gly Gly Ala Met Cys Trp Arg Arg Arg Arg
20 245 250 255

Ala Lys Pro Ser Glu Ser Arg His Pro Gly Pro Gly Ser Phe Gly
260 265 270

Arg Gly Gly Ser Leu Gly Leu Gly Gly Gly Gly Met Gly Pro
275 280 285

25 Arg Glu Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly
290 295 300

Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys Val Ser
305 310 315

Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro
30 320 325 330

Gln Ser Pro Pro Asn Ile Tyr Tyr Thr Ser Ile Ser Val Leu Glu
335 340 345

Trp Pro Ile Leu His Thr Ile Gln Leu Phe Phe Met Arg Ser Lys
350 355 360

35 Cys Ser Arg Val Thr Thr Phe Leu Phe Pro Val Gln Val Ile Thr

| | | | |
|----|---|-----|-----|
| | 365 | 370 | 375 |
| | Thr Ser Thr Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn | | |
| | 380 | 385 | 390 |
| | Pro Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg Ile | | |
| 5 | 395 | 400 | 405 |
| | Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr Ala Leu Phe | | |
| | 410 | 415 | 420 |
| | Val Leu Val Leu Ile Leu Leu Gly Arg Leu Asn Met His Gln | | |
| | 425 | 430 | 435 |
| 10 | Thr Thr Leu Leu Arg Gln Arg Ala Ser Val Glu Ala Glu Ala Gly | | |
| | 440 | 445 | 450 |
| | Gln His Gly Pro Leu | | |
| | 455 | | |

(2) INFORMATION FOR SEQ ID NO:3:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2380 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GNCTAGAAC TAGTGGATCC CCCCGGGCTG CAGGAATTCC GACGGCCCCT 50
GGAAGGGCTC TGGTGGGGCT GAGCGCTCTG CCGCGGGGGC GCAGGGCACAG 100
CAGGAAGCAG GTCCCGGTGG GCGCTGGGGG CATCAGCTAC CGGGGTGGTC 150
CGGGCTGAAG AGCCAGGCAG CCAAGGCAGC CACCCCGGGG GGTGGGCGAC 200
25 TTTGGGGGAG TTGGTGCCCC GCCCCCCCAGG CCTTGGCGGG GTCATGGGGC 250
CCCCCCATTC TGGGCCGGGG GGCGTGCGAG TCAGGGCCCT GCTGCTGCTG 300
GGGGTTTTGG GGCTGGTGTC TGGGCTCAGC CTGGAGCCTG TCTACTGGAA 350
CTCGGCGAAT AAGAGGTTCC AGGCAGAGGG TGGTTATGTG CTGTACCCCTC 400
AGATCAGGGGA CCGGCTAGAC CTGCTCTGCC CCCGGGCCCG GCCTCCTGGC 450
30 CCTCACTCCT CTCCTAATTA TGAGTTCTAC AAGCTGTACC TGGTAGGGGG 500
TGCTCAGGGC CGGCAGCTGTG AGGCACCCCC TGCCCCAAC CTCCTCTCA 550

CTTGTGATCG CCCAGACCTG GATCTCCGCT TCACCATCAA GTTCCAGGAG 600
TATAGCCCTA ATCTCTGGGG CCACGAGTTC CGCTCGCACC ACGATTACTA 650
CATCATTGCC ACATCGGATG GGACCCGGGA GGGCCTGGAG AGCCTGCAGG 700
GAGGTGTGTG CCTAACCAAGA GGCATGAAGG TGCTTCTCCG AGTGGGACAA 750
5 AGTCCCCGAG GAGGGGCTGT CCCCCGAAAA CCTGTGTCTG AAATGCCAT 800
GGAAAGAGAC CGAGGGGCAG CCCACAGCCT GGAGCCTGGG AAGGAGAAC 850
TGCCAGGTGA CCCCACCAGC AATGCAACCT CCCGGGGTGC TGAAGGCC 900
CTGCCCCCTC CCAGCATGCC TGCAGTGGCT GGGCAGCAG GGGGCTGGC 950
GCTGCTCTTG CTGGCGTGG CAGGGGCTGG GGGTGCCATG TGTTGGCGGA 1000
10 GACGGCGGGC CAAGCCTTCG GAGAGTCGCC ACCCTGGTCC TGGCTCCTTC 1050
GGGAGGGGAG GGTCTCTGGG CCTGGGGGGT GGAGGTGGGA TGGGACCTCG 1100
GGAGGCTGAG CCTGGGGAGC TAGGGATAGC TCTGCGGGGT GGCGGGGCTG 1150
CAGATCCCCC CTTCTGCC CACTATGAGA AGGTGAGTGG TGACTATGGG 1200
CATCCTGTGT ATATCGTGCA GGATGGGCC CCCAGAGCC CTCCAAACAT 1250
15 CTACTACAAG GTATGAGGGC TCCTCTCACG TGGCTATCCT GAATCCAGCC 1300
CTTCTTGGGG TGCTCCTCCA GTTTAATTCC TGGTTTGAGG GACACCTCTA 1350
ACATCTCGGC CCCCTGTGCC CCCCCAGCCC CTTCACTCCT CCCGGCTGCT 1400
GTCCTCGTCT CCACTTTAG GATTCCCTAG GATTCCCCT GCCCCACTTC 1450
CTGCCCTCCC GTTTGGCCAT GGGTCCCCCC CTCTGTCTCA GTGTCCCTGG 1500
20 ATCCTTTTC CTTGGGGAGG GGCACAGGCT CAGCCTCCTC TCTGACCATG 1550
ACCCAGGCAT CCTTGTCCCC CTCACCCACC CAGAGCTAGG GGCAGGAACA 1600
GCCCACCTT TGGTTGGCAC CGCCTTCTTT CTGCCTCTCA CTGGTTTCT 1650
CTTCTCTATC TCTTATTCTT TCCCTCTCTT CCGTCTCTAG GTCTGTTCTT 1700
CTTCCCTAGC ATCCTCCTCC CCACATCTCC TTTCACCCCTC TTGGCTTCTT 1750
25 ATCCTGTGCC TCTCCCCTCT CCTGGGTGGG GGCATCAAAG CATTTCTCCC 1800
CTTAGCTTTC AGCCCCCCTT CTGACCTCTC ATACCAACCA CTCCCCCTCAG 1850

TCTGCCAAAA ATGGGGCCT TATGGGAAG GCTCTGACAC TCCACCCAG 1900
CTCAGGCCAT GGGCAGCAGG GCTCCATTCT CTGGCCTGGC CCAGGCCTCT 1950
ACATAACTTAC TCCAGCCATT TGGGGTGGTT GGGTCATGAC AGCTACCAG 2000
AGAAGAAGTG TCCC GTTTG TCCAGTGGCC AATAGCAAGA TATGAACCGG 2050
5 TCGGGACATG TATGGACTTG GTCTGATGCT GAATGGGCCA CTTGGGACCG 2100
GAAGTGACTT GCTCCAGACA AGAGGTGACC AGGCCCGGAC AGAAATGGCC 2150
TGGGAAGTAG CAGAACAGT GCAGCAGGAA CTGGAAAGTGC CTTCATCCAG 2200
GACAGGAAGT AGCACTTCTG AAACAGGAAG TGGTCTGGCT GGAACCTCAA 2250
GTGGCTTAGT CTGGGGATC AGGAGGTGGG AGGTGGATGG TTCTTATTCT 2300
10 GTGGAGAAGA AGGGCGGGAA GAACTTCCTT TCAGGAGGAA GCTGGAACCTT 2350
ACTGACTGTA AGAGGTTAGA GGTGGACCGA 2380

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
15 (A) LENGTH: 340 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
1 5 10 15
20 Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
20 25 30
Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
35 40 45
25 Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
50 55 60
Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn
65 70 75
Tyr Glu Phe Tyr Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg
80 85 90
30 Arg Cys Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
95 100 105

| | | | |
|----|---|-----|-----|
| | Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln Glu Tyr | | |
| | 110 | 115 | 120 |
| | Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser His His Asp Tyr | | |
| | 125 | 130 | 135 |
| 5 | Tyr Ile Ile Ala Thr Ser Asp Gly Thr Arg Glu Gly Leu Glu Ser | | |
| | 140 | 145 | 150 |
| | Leu Gln Gly Gly Val Cys Leu Thr Arg Gly Met Lys Val Leu Leu | | |
| | 155 | 160 | 165 |
| 10 | Arg Val Gly Gln Ser Pro Arg Gly Gly Ala Val Pro Arg Lys Pro | | |
| | 170 | 175 | 180 |
| | Val Ser Glu Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser | | |
| | 185 | 190 | 195 |
| | Leu Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser Asn | | |
| | 200 | 205 | 210 |
| 15 | Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro Pro Ser Met | | |
| | 215 | 220 | 225 |
| | Pro Ala Val Ala Gly Ala Ala Gly Gly Leu Ala Leu Leu Leu Leu | | |
| | 230 | 235 | 240 |
| 20 | Gly Val Ala Gly Ala Gly Gly Ala Met Cys Trp Arg Arg Arg Arg | | |
| | 245 | 250 | 255 |
| | Ala Lys Pro Ser Glu Ser Arg His Pro Gly Pro Gly Ser Phe Gly | | |
| | 260 | 265 | 270 |
| | Arg Gly Gly Ser Leu Gly Leu Gly Gly Gly Gly Met Gly Pro | | |
| | 275 | 280 | 285 |
| 25 | Arg Glu Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly | | |
| | 290 | 295 | 300 |
| | Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys Val Ser | | |
| | 305 | 310 | 315 |
| 30 | Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro | | |
| | 320 | 325 | 330 |
| | Gln Ser Pro Pro Asn Ile Tyr Tyr Lys Val | | |
| | 335 | 340 | |

(2) INFORMATION FOR SEQ ID NO:5:

- (i) SEQUENCE CHARACTERISTICS:
- 35 (A) LENGTH: 454 base pairs

- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

5 GCCGACGCTG TGAGGCACCC CCTGCCCAA ACCTCCTTCT CACTTGTGAT 50
CGCCCAGACC TGGATCTCCG CTTCACCATC AAGTTCCAGG AGTATAGCCC 100
TAATCTCTGG GGCCACGAGT TCCGCTCGCA CCACGATTAC TACATCATTG 150
CCACATCGGA TGGGACCCGG GAGGCCTGGG AGAGCCTGCA GGGAAAGTGTG 200
TGCCTAACCA GAGGCATGAA GGTGCTTCTC CGAGTNGGAC AAAGTCCCGA 250
10 GGAGGGGCTG TCCCCCGAAA ACCTGTGTCT GAAATGCCCA TGGAAAGAGA 300
CCGAGGGGCA GCCCACAGCC TGGGAGCCTG GGGAAAGGAGA ACCTGCCAGG 350
TGACCCCCACC AGCAATNCAA CCTTCCGGGG TTGCTTGAAG GGCCCCTTGA 400
CCCTTTCCCA GCATTGCNTG CANTTGGTTN GGGCAGCAN GGGGGNGTTT 450
TGGC 454

15 (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

30 GGACAAAGTC CCGAGGGAGGG GCTGTCCCCC GAAAACCTGT GTCTGAAATG 50

CCCATGGAAA 60

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 60 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGGTTCTCC TTCCCCAGGC TCCCAGGCTG TGGGCTGCC CTCGGTCTCT 50

10 TTCCATGGGC 60



SEQUENCE LISTING

GENERAL INFORMATION:

SUB
JB2

(i) APPLICANT: Caras, Ingrid W
(ii) TITLE OF INVENTION: A2-1 Neurotrophic Factor

(iii) NUMBER OF SEQUENCES: 8

10 (iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 460 Point San Bruno Blvd
(C) CITY: South San Francisco
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080

15 (v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: WinPatin (Genentech)

20 (vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: 08/635130
(B) FILING DATE: 19-Mar-1996
(C) CLASSIFICATION:

25 (viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Torchia, PhD., Timothy E.
(B) REGISTRATION NUMBER: 36,700
(C) REFERENCE/DOCKET NUMBER: P1001

30 (ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 415/225-8674
(B) TELEFAX: 415/952-9881
(C) TELEX: 910/371-7168

35 (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1877 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Double
(D) TOPOLOGY: Linear

40 (ix) FEATURE:

(A) NAME/KEY: Extra Cellular Domain
(B) LOCATION: 244-899
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(ix) FEATURE:
(A) NAME/KEY: Transmembrane Domain
(B) LOCATION: 901-978
(C) IDENTIFICATION METHOD:
5 (D) OTHER INFORMATION:

(ix) FEATURE:
(A) NAME/KEY: signal peptide
(B) LOCATION: 244-321
(C) IDENTIFICATION METHOD:
10 (D) OTHER INFORMATION:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GNTCTAGAAN TAGTGGATCC CCCCAGGCTG CAGGAATTCC GACGGCCCC 50

GGAAGGGCTC TGGTGGGGCT GAGCGCTCTG CCGCGGGGGC GCGGGCACAG 100

CAGGAAGCAG GTCCCGGTGG GCGCTGGGGG CATCAGCTAC CGGGGTGGTC 150

15 CGGGCTGAAG AGCCAGGCAG CCAAGGCAGC CACCCCGGGG GGTGGGCAG 200

TTTGGGGGAG TTGGTGCCCC GCCCCCCCAGG CCTGGCGGG GTC ATG 246

Met
1

GGG CCC CCC CAT TCT GGG CCG GGG GGC GTG CGA GTC GGG 285

20 Gly Pro Pro His Ser Gly Pro Gly Val Arg Val Gly
5 10

GCC CTG CTG CTG CTG GGG GTT TTG GGG CTG GTG TCT GGG 324

Ala Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly
15 20 25



25 CTC AGC CTG GAG CCT GTC TAC TGG AAC TCG GCG AAT AAG 363
Leu Ser Leu Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys
30 35 40

AGG TTC CAG GCA GAG GGT GGT TAT GTG CTG TAC CCT CAG 402
Arg Phe Gln Ala Glu Gly Gly Tyr Val Leu Tyr Pro Gln
30 45 50

ATC GGG GAC CGG CTA GAC CTG CTC TGC CCC CGG GCC CGG 441
Ile Gly Asp Arg Leu Asp Leu Leu Cys Pro Arg Ala Arg
55 60 65

35 CCT CCT GGC CCT CAC TCC TCT CCT AAT TAT GAG TTC TAC 480
Pro Pro Gly Pro His Ser Ser Pro Asn Tyr Glu Phe Tyr
70 75

AAG CTG TAC CTG GTA GGG GGT GCT CAG GGC CGG CGC TGT 519

Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg Arg Cys
 80 85 90

GAG GCA CCC CCT GCC CCA AAC CTC CTT CTC ACT TGT GAT 558
 Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
 5 95 100 105

CGC CCA GAC CTG GAT CTC CGC TTC ACC ATC AAG TTC CAG 597
 Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln
 110 115

GAG TAT AGC CCT AAT CTC TGG GGC CAC GAG TTC CGC TCG 636
 10 Glu Tyr Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser
 120 125 130

CAC CAC GAT TAC TAC ATC ATT GCC ACA TCG GAT GGG ACC 675
 His His Asp Tyr Tyr Ile Ile Ala Thr Ser Asp Gly Thr
 135 140

15 CGG GAG GGC CTG GAG AGC CTG CAG GGA GGT GTG TGC CTA 714
 Arg Glu Gly Leu Glu Ser Leu Gln Gly Gly Val Cys Leu
 145 150 155

ACC AGA GGC ATG AAG GTG CTT CTC CGA GTG GGA CAA AGT 753
 Thr Arg Gly Met Lys Val Leu Leu Arg Val Gly Gln Ser
 20 160 165 170

CCC CGA GGA GGG GCT GTC CCC CGA AAA CCT GTG TCT GAA 792
 Pro Arg Gly Gly Ala Val Pro Arg Lys Pro Val Ser Glu
 175 180

ATG CCC ATG GAA AGA GAC CGA GGG GCA GCC CAC AGC CTG 831
 25 Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser Leu
 185 190 195

GAG CCT GGG AAG GAG AAC CTG CCA GGT GAC CCC ACC AGC 870
 Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser
 200 205

30 AAT GCA ACC TCC CGG GGT GCT GAA GGC [CCC] CTG CCC CCT 909
 Asn Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro
 210 215 220

CCC AGC ATG CCT GCA GTG GCT GGG GCA GCA GGG GGG CTG 948
 Pro Ser Met Pro Ala Val Ala Gly Ala Ala Gly Gly Leu
 35 225 230 235

GCG CTG CTC TTG CTG GGC GTG GCA GGG GCT GGG GGT GCC 987
 Ala Leu Leu Leu Leu Gly Val Ala Gly Ala Gly Gly Ala
 240 245

ATG TGT TGG CGG AGA CGG CGG GCC AAG CCT TCG GAG AGT 1026

Met Cys Trp Arg Arg Arg Arg Ala Lys Pro Ser Glu Ser
 250 255 260

CGC CAC CCT GGT CCT GGC TCC TTC GGG AGG GGA GGG TCT 1065
 Arg His Pro Gly Pro Gly Ser Phe Gly Arg Gly Gly Ser
 5 265 270

CTG GGC CTG GGG GGT GGA GGT ATG GGA CCT CGG GAG 1104
 Leu Gly Leu Gly Gly Gly Met Gly Pro Arg Glu
 275 280 285

GCT GAG CCT GGG GAG CTA GGG ATA GCT CTG CGG GGT GGC 1143
 10 Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly
 290 295 300

GGG GCT GCA GAT CCC CCC TTC TGC CCC CAC TAT GAG AAG 1182
 Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys
 305 310

15 GTG AGT GGT GAC TAT GGG CAT CCT GTG TAT ATC GTG CAG 1221
 Val Ser Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln
 315 320 325

GAT GGG CCC CCC CAG AGC CCT CCA AAC ATC TAC TAG ACA 1260
 Asp Gly Pro Pro Gln Ser Pro Pro Asn Ile Tyr Tyr Thr
 20 330 335

TCG ATT TCT GTG TTG GAG TGG CCC ATA TTG CAT ACG ATA 1299
 Ser Ile Ser Val Leu Glu Trp Pro Ile Leu His Thr Ile
 340 345 350

CAA CTG TTT TTC ATG CGA TCC AAG TGG TCC CGT GTC ACT 1338
 25 Gln Leu Phe Phe Met Arg Ser Lys Cys Ser Arg Val Thr
 355 360 365

ACA TTC TTA TTT CCT GTG CAA GTT ATT ACG ACA TCG ACT 1377
 Thr Phe Leu Phe Pro Val Gln Val Ile Thr Thr Ser Thr
 370 375

30 TGC CGG ATG ACT TCA TTT AGC TTT ACC ACC CTG AAC CCA 1416
 Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn Pro
 380 385 390

TCC ATG CAG GCC TGC AGA GCA CAG ATG GGG GAA TTC CGA 1455
 Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg
 35 395 400

ATC AGA TGG TGT TTC TGG GGG GAC AGG ATC CTG GGT ACG 1494
 Ile Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr
 405 410 415

GCT CTG TTT GTG CTT GTG CTT ATT CTT CTT CTT GGG AGG 1533

Ala Leu Phe Val Leu Val Leu Ile Leu Leu Leu Gly Arg
420 425 430

CTG AAT ATG CAT CAG ACG ACA CTG CTC CGG CAA CGG GCC 1572
Leu Asn Met His Gln Thr Thr Leu Leu Arg Gln Arg Ala
5 435 440

AGT GTG GAG GCG GAA GCC GGC CAG CAT GGT CCC CTG TG 1610
Ser Val Glu Ala Glu Ala Gly Gln His Gly Pro Leu
445 450 455

ATAGGATTGA AAGAGCTACT GAGAATAGGG GGCTTCTCAA TGAGAGAGCG 1660

10 GAGGCTGCTG TTATCATGGG AACCAAGGCAG ATCAATCATC CCTGGCAGGT 1710
CAGGCAGGAA GTTACTTAGC TTCTCCTTCA CCTTCTTCCC ACAGAATTAA 1760
TTATAGGCTT GTTCCAAGTT GTAGTGTGTG ATCAGATTG TGCTGCCTGT 1810
CAGCTCTGTG CTACCTGGCA GTTCCCCCTCA TGGAAATTGAA TATCAAGCTT 1860
ATCGATAACCG TCGACCT 1877

15 (2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 455 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
1 5 10 15

Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
20 25 30

25 Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
35 40 45

Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
50 55 60

30 Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn
65 70 75

Tyr Glu Phe Tyr Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg
80 85 90

Arg Cys Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
95 100 105

Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln Glu Tyr
110 115 120

Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser His His Asp Tyr
125 130 135

5 Tyr Ile Ile Ala Thr Ser Asp Gly Thr Arg Glu Gly Leu Glu Ser
140 145 150

Leu Gln Gly Gly Val Cys Leu Thr Arg Gly Met Lys Val Leu Leu
155 160 165

Arg Val Gly Gln Ser Pro Arg Gly Gly Ala Val Pro Arg Lys Pro
10 170 175 180

Val Ser Glu Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser
185 190 195

Leu Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser Asn
200 205 210

15 Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro Pro Ser Met
215 220 225

Pro Ala Val Ala Gly Ala Ala Gly Gly Leu Ala Leu Leu Leu
230 235 240

Gly Val Ala Gly Ala Gly Gly Ala Met Cys Trp Arg Arg Arg Arg
20 245 250 255

Ala Lys Pro Ser Glu Ser Arg His Pro Gly Pro Gly Ser Phe Gly
260 265 270

Arg Gly Gly Ser Leu Gly Leu Gly Gly Gly Gly Met Gly Pro
275 280 285

25 Arg Glu Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly
290 295 300

Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys Val Ser
305 310 315

Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro
320 325 330

Gln Ser Pro Pro Asn Ile Tyr Tyr Ser Ile Ser Val Leu Glu
335 340 345

Tyr Pro Ile Leu His Thr Ile Gln Leu Phe Phe Met Arg Ser Lys
350 355 360

35 Cys Ser Arg Val Thr Thr Phe Leu Phe Pro Val Gln Val Ile Thr

365 370 375

Thr Ser Thr Cys Arg Met Thr Ser Phe Ser Phe Thr Thr Leu Asn
380 385 390

5 Pro Ser Met Gln Ala Cys Arg Ala Gln Met Gly Glu Phe Arg Ile
395 400 405

Arg Trp Cys Phe Trp Gly Asp Arg Ile Leu Gly Thr Ala Leu Phe
410 415 420

Val Leu Val Leu Ile Leu Leu Gly Arg Leu Asn Met His Gln
425 430 435

10 Thr Thr Leu Leu Arg Gln Arg Ala Ser Val Glu Ala Glu Ala Gly
440 445 450

Gln His Gly Pro Leu
455

(2) INFORMATION FOR SEQ ID NO:3:

15 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2380 base pairs
(B) TYPE: Nucleic Acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GNCTAGAAC TAGTGGATCC CCCCGGGCTG CAGGAATTCC GACGGCCCCT 50
GGAAGGGCTC TGGTGGGGCT GAGCGCTCTG CCGCGGGGGC GCAGGGCACAG 100
CAGGAAGCAG GTCCCGGTGG GCGCTGGGGG CATCAGCTAC CGGGGTGGTC 150
CGGGCTGAAG AGCCAGGCAG CCAAGGCAGC CACCCCGGGG GGTGGGCGAC 200
25 TTTGGGGGAG TTGGTGCCCC GCCCCCCCAGG CCTTGGCGGG GTCATGGGC 250
CCCCCCATTTC TGGGCCGGGG GGCGTGCGAG TCGGGGCCCT GCTGCTGCTG 300
GGGGTTTTGG GGCTGGTGTC TGGGCTCAGC CTGGAGCCTG TCTACTGGAA 350
CTCGGCGAAT AAGAGGTTCC AGGCAGAGGG TGGTTATGTG CTGTACCCCTC 400
AGATCGGGGA CGGGCTAGAC CTGCTCTGCC CCCGGGCCCG GCCTCCTGGC 450
30 CCTCACTCCT CTCCTAATTA TGAGTTCTAC AAGCTGTACC TGGTAGGGGG 500
TGCTCAGGGC CGGCAGCTGTG AGGCACCCCC TGCCCCAAC CTCCTTCTCA 550

CTTGTGATCG CCCAGACCTG GATCTCCGCT TCACCATCAA GTTCCAGGAG 600
TATAGCCCTA ATCTCTGGGG CCACGAGTTC CGCTCGCACC ACGATTACTA 650
CATCATTGCC ACATCGGATG GGACCCGGGA GGGCCTGGAG AGCCTGCAGG 700
GAGGTGTGTG CCTAACCAAGA GGCATGAAGG TGCTTCTCCG AGTGGGACAA 750
5 AGTCCCCGAG GAGGGGCTGT CCCCCGAAAA CCTGTGTCTG AAATGCCAT 800
GGAAAGAGAC CGAGGGGCAG CCCACAGCCT GGAGCCTGGG AAGGAGAAC 850
TGCCAGGTGA CCCCACCAGC AATGCAACCT CCCGGGGTGC TGAAGGCC 900
CTGCCCCCTC CCAGCATGCC TGCAGTGGCT GGGCAGCAG GGGGGCTGCC 950
GCTGCTCTTG CTGGCGTGG CAGGGGCTGG GGGTGCCATG TGTTGGCGGA 1000
10 GACGGCGGGC CAAGCCTTCG GAGAGTCGCC ACCCTGGTCC TGGCTCCTTC 1050
GGGAGGGGAG GGTCTCTGGG CCTGGGGGGT GGAGGTGGGA TGGGACCTCG 1100
GGAGGCTGAG CCTGGGGAGC TAGGGATAGC TCTGCGGGGT GGCAGGGCTG 1150
CAGATCCCCC CTTCTGCC CACTATGAGA AGGTGAGTGG TGACTATGGG 1200
CATCCTGTGT ATATCGTGCA GGATGGGCC CCCCAGAGCC CTCCAAACAT 1250
15 CTACTACAAG GTATGAGGGC TCCTCTCACG TGGCTATCCT GAATCCAGCC 1300
CTTCTTGGGG TGCTCCTCCA GTTTAATTCC TGGTTTGAGG GACACCTCTA 1350
ACATCTCGGC CCCCTGTGCC CCCCCAGCCC CTTCACTCCT CCCGGCTGCT 1400
GTCCTCGTCT CCACTTTAG GATTCCCTAG GATTCCCCT GCCCCACTTC 1450
CTGCCCTCCC GTTGGCCAT GGGTGCCCCC CTCTGTCTCA GTGTCCCTGG 1500
20 ATCCTTTTTC CTTGGGGAGG GGCACAGGCT CAGCCTCCTC TCTGACCATG 1550
ACCCAGGCAT CCTTGCCCCC CTCACCCACC CAGAGCTAGG GGCAGGAACA 1600
GCCACCTTT TGGTTGGCAC CGCCTTCTTT CTGCCTCTCA CTGGTTTCT 1650
CTTCTCTATC TCTTATTCTT TCCCTCTCTT CCGTCTCTAG GTCTGTTCTT 1700
CTTCCCTAGC ATCCTCCTCC CCACATCTCC TTTCACCCCTC TTGGCTTCTT 1750
25 ATCCTGTGCC TCTCCCCTCT CCTGGGTGGG GGCATCAAAG CATTCTCCC 1800
CTTAGCTTTC AGCCCCCTT CTGACCTCTC ATACCAACCA CTCCCCCTCAG 1850

TCTGCCAAAA ATGGGGGCCT TATGGGAAG GCTCTGACAC TCCACCCAG 1900
CTCAGGCCAT GGGCAGCAGG GCTCCATTCT CTGGCCTGGC CCAGGCCTCT 1950
ACATAACTTAC TCCAGCCATT TGGGGTGGTT GGGTCATGAC AGCTACCATG 2000
AGAAGAAGTG TCCC GTTTG TCCAGTGGCC AATAGCAAGA TATGAACCGG 2050
5 TCGGGACATG TATGGACTTG GTCTGATGCT GAATGGGCCA CTTGGGACCG 2100
GAAGTGACTT GCTCCAGACA AGAGGTGACC AGGCCCGGAC AGAAATGGCC 2150
TGGGAAGTAG CAGAACAGT GCAGCAGGAA CTGGAAAGTGC CTTCATCCAG 2200
GACAGGAAGT AGCACTTCTG AAACAGGAAG TGGTCTGGCT GGAACCTCAA 2250
GTGGCTTAGT CTGGGGATC AGGAGGTGGG AGGTGGATGG TTCTTATTCT 2300
10 GTGGAGAAGA AGGGCGGGAA GAACTTCCTT TCAGGAGGAA GCTGGAACCTT 2350
ACTGACTGTA AGAGGTTAGA GGTGGACCGA 2380

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

15 (A) LENGTH: 340 amino acids
(B) TYPE: Amino Acid
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Pro Pro His Ser Gly Pro Gly Gly Val Arg Val Gly Ala
1 5 10 15
20 Leu Leu Leu Leu Gly Val Leu Gly Leu Val Ser Gly Leu Ser Leu
20 25 30
Glu Pro Val Tyr Trp Asn Ser Ala Asn Lys Arg Phe Gln Ala Glu
35 40 45
25 Gly Gly Tyr Val Leu Tyr Pro Gln Ile Gly Asp Arg Leu Asp Leu
50 55 60
Leu Cys Pro Arg Ala Arg Pro Pro Gly Pro His Ser Ser Pro Asn
65 70 75
Tyr Glu Phe Tyr Lys Leu Tyr Leu Val Gly Gly Ala Gln Gly Arg
80 85 90
30 Arg Cys Glu Ala Pro Pro Ala Pro Asn Leu Leu Leu Thr Cys Asp
95 100 105

| | | | |
|----|---|-----|-----|
| | Arg Pro Asp Leu Asp Leu Arg Phe Thr Ile Lys Phe Gln Glu Tyr | | |
| | 110 | 115 | 120 |
| | Ser Pro Asn Leu Trp Gly His Glu Phe Arg Ser His His Asp Tyr | | |
| | 125 | 130 | 135 |
| 5 | Tyr Ile Ile Ala Thr Ser Asp Gly Thr Arg Glu Gly Leu Glu Ser | | |
| | 140 | 145 | 150 |
| | Leu Gln Gly Gly Val Cys Leu Thr Arg Gly Met Lys Val Leu Leu | | |
| | 155 | 160 | 165 |
| 10 | Arg Val Gly Gln Ser Pro Arg Gly Gly Ala Val Pro Arg Lys Pro | | |
| | 170 | 175 | 180 |
| | Val Ser Glu Met Pro Met Glu Arg Asp Arg Gly Ala Ala His Ser | | |
| | 185 | 190 | 195 |
| | Leu Glu Pro Gly Lys Glu Asn Leu Pro Gly Asp Pro Thr Ser Asn | | |
| | 200 | 205 | 210 |
| 15 | Ala Thr Ser Arg Gly Ala Glu Gly Pro Leu Pro Pro Pro Ser Met | | |
| | 215 | 220 | 225 |
| | Pro Ala Val Ala Gly Ala Ala Gly Gly Leu Ala Leu Leu Leu Leu | | |
| | 230 | 235 | 240 |
| 20 | Gly Val Ala Gly Ala Gly Gly Ala Met Cys Trp Arg Arg Arg Arg | | |
| | 245 | 250 | 255 |
| | Ala Lys Pro Ser Glu Ser Arg His Pro Gly Pro Gly Ser Phe Gly | | |
| | 260 | 265 | 270 |
| | Arg Gly Gly Ser Leu Gly Leu Gly Gly Gly Gly Met Gly Pro | | |
| | 275 | 280 | 285 |
| 25 | Arg Glu Ala Glu Pro Gly Glu Leu Gly Ile Ala Leu Arg Gly Gly | | |
| | 290 | 295 | 300 |
| | Gly Ala Ala Asp Pro Pro Phe Cys Pro His Tyr Glu Lys Val Ser | | |
| | 305 | 310 | 315 |
| 30 | Gly Asp Tyr Gly His Pro Val Tyr Ile Val Gln Asp Gly Pro Pro | | |
| | 320 | 325 | 330 |
| | Gln Ser Pro Pro Asn Ile Tyr Tyr Lys Val | | |
| | 335 | 340 | |

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 454 base pairs

- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

5 GCCGACGCTG TGAGGCACCC CCTGCCCAA ACCTCCTTCT CACTTGAT 50
CGCCCAGACC TGGATCTCCG CTTCACCATC AAGTTCCAGG AGTATAGCCC 100
TAATCTCTGG GGCCACGAGT TCCGCTCGCA CCACGATTAC TACATCATTG 150
CCACATCGGA TGGGACCCGG GAGGCCTGGG AGAGCCTGCA GGGAAAGTGTG 200
TGCCTAACCA GAGGCATGAA GGTGCTTCTC CGAGTNGGAC AAAGTCCCGA 250
10 GGAGGGGCTG TCCCCCGAAA ACCTGTGTCT GAAATGCCCA TGGAAAGAGA 300
CCGAGGGGCA GCCCACAGCC TGGGAGCCTG GGGAAAGGAGA ACCTGCCAGG 350
TGACCCCCACC AGCAATNCAA CCTTCCGGGG TTGCTTGAAG GGCCCCTTGA 400
CCCTTTCCCA GCATTGCNTG CANTTGGTTN GGGCAGCAN GGGGGNGTTT 450
TGGC 454

15 (2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 amino acids
 - (B) TYPE: Amino Acid
 - (D) TOPOLOGY: Linear

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Asp Lys Thr His Thr Cys Pro Pro Cys Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 60 base pairs
 - (B) TYPE: Nucleic Acid
 - (C) STRANDEDNESS: Single
 - (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

30 GGACAAAGTC CCGAGGGAGGG GCTGTCCCCC GAAAACCTGT GTCTGAAATG 50

CCCATGGAAA 60

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 60 base pairs
- (B) TYPE: Nucleic Acid
- (C) STRANDEDNESS: Single
- (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGGTTCTCC TTCCCCAGGC TCCCAGGCTG TGGGCTGCC CTCGGTCTCT 50

10 TTCCATGGGC 60

FILE DKB2:[MANDL]HS00616.EMEST1;1
ID HS006163 standard; RNA; EST; 454 BP.
XX
AC H10006;
XX
NI g874828
XX
DT 02-JUL-1995 (Rel. 44, Created)
DT 02-JUL-1995 (Rel. 44, Last updated, Version 1)
XX
DE ym02e11.r1 Homo sapiens cDNA clone 46610 5' similar to contains Alu
DE repetitive element;. .
XX
KW EST.
XX
OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-454
RA Hillier L., Clark N., Dubuque T., Elliston K., Hawkins M.,
RA Holman M., Hultman M., Kucaba T., Le M., Lennon G., Marra M.,
RA Parsons J., Rifkin L., Rohlfing T., Soares M., Tan F.,
RA Trevaskis E., Waterston R., Williamson A., Wohldmann P., Wilson R.;
RT "The WashU-Merck EST Project";
RL Unpublished.
XX
CC GDB: G00-419-151 Contact: Wilson RK WashU-Merck EST Project
CC Washington University School of Medicine 4444 Forest Park Parkway,
CC Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810
CC Email: est@watson.wustl.edu High quality sequence stops: 313
CC Source: IMAGE Consortium, LLNL This clone is available royalty-free
CC through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov)
CC for further information. NCBI gi: 874828

XX
FH Key Location/Qualifiers
FH
FT source 1..454
FT /organism="Homo sapiens"
FT /clone="46610"
FT /note="human"
XX
SQ Sequence 454 BP; 93 A; 137 C; 133 G; 84 T; 7 other;

HS006163 Length: 454 August 19, 1997 11:01 Type: N Check: 5608 ..

1 GCCGACGCTG TGAGGCACCC CCTGCCCAA ACCTCCTTCT CACTTGTGAT
51 CGCCCAGACC TGGATCTCCG CTTCACCATC AAGTTCCAGG AGTATAGCCC
101 TAATCTCTGG GGCCACGAGT TCCGCTCGCA CCACGATTAC TACATCATTG
151 CCACATCGGA TGGGACCCGG GAGGCCTGGG AGAGCCTGCA GGGAAAGTGTG
201 TGCCTAACCA GAGGCATGAA GGTGCTTCTC CGAGTNGGAC AAAGTCCCGA

251 GGAGGGGCTG TCCC[REDACTED]AAA ACCTGTGTCT GAAATGCCCA[REDACTED]GGAAAGAGA
301 CCGAGGGGCA GCCCACAGCC TGGGAGCCTG GGGAAAGGAGA ACCTGCCAGG
351 TGACCCCCACC AGCAATNCAA CCTTCCGGGG TTGCTTGAAG GGCCCCTTGA
401 CCCTTTCCCA GCATTGCNTG CANTTGGTTN GGGGCAGCAN GGGGGNGTTT
451 TGGC

ile DKB2:[MANDL]U57001.EMHUM2;1
D HSU57001 standard; RNA; HUM; 1708 BP.
X
.C U57001;
X
II g1469781
XX
DT 31-JUL-1996 (Rel. 48, Created)
DT 31-JUL-1996 (Rel. 48, Last updated, Version 1)
XX DE Human ligand for eph-related receptor tyrosine kinases (EPLG8)
DE mRNA, complete cds.
XX
KW
XX OS Homo sapiens (human)
OC Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;
OC Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX
RN [1]
RP 1-1708
RA Cerretti D.P.;
RT "LERK-8, a ligand for the eph-related receptor tyrosine kinases";
RL Unpublished.
XX
RN [2]

RP 1-1708
RA Cerretti D.P.;
RT ; Submitted (29-APR-1996) to the EMBL/GenBank/DDBJ databases.
RL Douglas P. Cerretti, Molecular Biology, Immunex Corp., 51
RL University St., Seattle, WA 98101, USA
XX DR SPTREMBL; Q15768; Q15768.
XX FH Key Location/Qualifiers
FH FT source 1. .1708
FT /organism="Homo sapiens"
FT /clone_lib="Clontech number HL3003a"
FT FT sig_peptide 398. .466
FT /gene="EPLG8"
FT CDS 398. .1420
FT /gene="EPLG8"
FT /note="LERK-8"
FT /codon_start=1
FT /product="ligand for eph-related receptor tyrosine kinase"
FT /db_xref="PID:g1469782"
FT /db_xref="SPTREMBL:Q15768"
FT /translation="MGPPHSGPGGVVGALLLGVLGLVSGLSLEPVYWNNSANKRFQ
GGYVLYPQIGDRLLCPRARPPGPSSPNYEFYKLYLVGGAQGRRCEAPPAPNLL
DRPDLDLRFTIKFQEYSPNLWGHEFRSHHDYYIATSDGTREGLESQGGVCLTRGM
FT

FT LLRVPRGGAVPRKPVSEMPMERDRGAAH PGKENLPGDPTSNTSRGAEGPLP
FT PSMPAVAGAAGGLALLLGAGAGGAMCWRRRRAKPSERSHPGPGSFGRGGSLGLGGGC
FT GMGPREAEPGELGIALRGGAADPPFCPHYEKVSGDYGHPVYIVQDGPPQSPPNIYYK
FT "
FT misc_feature 1070..1150
FT /gene="EPLG8"
FT /note="encodes transmembrane domain"
XX
SQ Sequence 1708 BP; 244 A; 563 C; 572 G; 329 T; 0 other;
U57001 Length: 1708 August 20, 1997 12:03 Type: N Check: 4485 ..
1 AATTCCGGCC CTTAGCCCCGC TGCCCTCAAT CCCAGCGAGG CTGGGGCTCC
51 GGCTCGGCGC CCCCTTCCTC GCTCCCTGGT CGGGCGCCCC ATGCCGCC
101 CGCCCCGGTCC CCGGCTCCCC CAGTCCCCCA CTTAGGCGGG CTCACAGATC
151 CCGGGGTGCT GGCGCGTGCG CCGGGGCGC GTAGGGCGCC TGCAGACGGC
201 CCCTGGAAGG GCTCTGGTGG GGCTGAGCGC TCTGCCGCGG GGGCGCGGGC
251 ACAGCAGGAA GCAGGTCCGC GTGGCGCTG GGGGCATCAG CTACCGGGGT
301 GGTCCGGGCT GAAGAGCCAG GCAGCCAAGG CAGCCACCCC GGGGGGTGGG
351 CGACTTTGGG GGAGTTGGTG CCCC GCCCCCCC CAGGCCTTGG CGGGGTCA
401 GGGCCCCCCC ATTCTGGGCC GGGGGCGTG CGAGTCGGGG CCTGCTGCT
451 GCTGGGGTT TTGGGGCTGG TGTCTGGCT CAGCCTGGAG CCTGTCTACT
501 GGAACCTCGGC GAATAAGAGG TTCCAGGCAG AGGGTGGTTA TGTGCTGTAC
551 CCTCAGATCG GGGACCGGCT AGACCTGCTC TGCCCCGGG CCCGGCCTCC
601 TGGCCCTCAC TCCTCTCCTA ATTATGAGTT CTACAAGCTG TACCTGGTAG
651 GGGGTGCTCA GGGCCGGCGC TGTGAGGCAC CCCCTGCCAA AACCTCCTT
701 CTCACTTGTG ATCGCCCAGA CCTGGATCTC CGCTTCACCA TCAAGTTCCA
751 GGAGTATAGC CCTAATCTCT GGGGCCACGA GTTCCGCTCG CACCACGATT
801 ACTACATCAT TGCCACATCG GATGGGACCC GGGAGGGCCT GGAGAGCCTG
851 CAGGGAGGTG TGTGCCTAAC CAGAGGCATG AAGGTGCTTC TCCGAGTGGG
901 ACAAAAGTCCC CGAGGGAGGGG CTGTCCCCCG AAAACCTGTG TCTGAAATGC
951 CCATGGAAAG AGACCGAGGG GCAGCCCACA GCCTGGAGCC TGGGAAGGAG
1001 AACCTGCCAG GTGACCCCAC CAGCAATGCA ACCTCCCGGG GTGCTGAAGG
1051 CCCCTGCC CCTCCCAGCA TGCCTGCAGT GGCTGGGGCA GCAGGGGGC

3101 TGGCGCTGCT CTTGGGC GTGGCAGGGG CTGGGGGT CATGTGTTGG
1151 CGGAGACGGC GGGCCAAGCC TTCGGAGAGT CGCCACCCCTG GTCCTGGCTC
1201 CTTCGGGAGG GGAGGGTCTC TGGGCCTGGG GGGTGGAGGT GGGATGGGAC
1251 CTCGGGAGGC TGAGCCTGGG GAGCTAGGGA TAGCTCTGCG GGGTGGCGGG

1301 GCTGCAGATC CCCCCCTCTG CCCCCACTAT GAGAAGGTGA GTGGTGACTA
1351 TGGGCATCCT GTGTATATCG TGCAGGATGG GCCCCCCCAG AGCCCTCCAA
1401 ACATCTACTA CAAGGTATGA GGGCTCCTCT CACGTGGCTA TCCTGAATCC
1451 AGCCCTTCTT GGGGTGCTCC TCCAGTTAA TTCCCTGGTTT GAGGGACACC
1501 TCTAACATCT CGGCCCTCTG TGCCCCCCCAC GCCCCTTCAC TCCTCCCGGC
1551 TGCTGTCCCTC GTCTCCACTT TTAGGATTCC TTAGGATTCC CACTGCCCA
1601 CTTCCTGCC C TCCCGTTGG CCATGGGTGC CCCCCTCTGT CTCAGTGTCC
1651 CTGGATCCTT TTTCCCTGGG GAGGGGCACA GGCTCAGCCT CCTCTCTGAC
1701 CATGCCGG